Appendix E

Methodological Notes

METHODOLOGICAL NOTES

This appendix provides the details of methods used for the evaluation of the Basic Pilot program. Included here are discussions of the statistical methods used in data analysis (Section A), establishment matching (Section B), cost estimates (Section C), and estimation of the work-authorization status of unresolved cases (Section D).

A. STATISTICAL METHODS USED IN DATA ANALYSIS

This section provides additional detail on the methods used in weighting, nonresponse adjustment, variance estimation, and statistical testing for the employer mail survey, the employer on-site data collection, and the employee in-person interviews.

1. BACKGROUND

Sample selection, nonresponse adjustment, and estimation for the employer survey and the employee interviews were based on probability sampling methods. Probability samples allow analysts to compute sampling weights and nonresponse adjustment weights, estimate the precision of sample estimates, and test for the statistical significance of study hypotheses. Nonresponse adjustment through weighting implies that, within adjustment cells, nonrespondents are similar to respondents in the characteristics or behavior reported in the survey. To the extent that this assumption is not correct, there may be some bias in the estimates.

2. WEIGHTING AND NONRESPONSE ADJUSTMENT

a. EMPLOYER SURVEY WEIGHTS

The pilot employer mail survey was a census of all eligible employers participating in the Basic Pilot during the study period. The non-pilot sample consisted of non-pilot establishments matched to the eligible pilot establishments, and no sampling per se was involved. Consequently, weighting was used only to adjust for nonresponse.[1] Since the pilot and non-pilot samples were identified using different methodologies, nonresponse weighting was done separately for the two groups.

The evaluation team examined three variables to define nonresponse adjustment cells: establishment size, sales volume, and Standard Industrial Classification (SIC) code. These three variables were available for respondents as well as nonrespondents and were likely to affect employer responses. Regression analysis was used to determine whether or not there was a significant relationship between response rates and each of these variables. Sales volume was found to be highly correlated with establishment size but was missing more often than establishment size. SIC code generated too many adjustment cells, and the cell sizes were too small. Ultimately, only establishment size was used for nonresponse adjustment for the mail survey and the on-site survey. The

[1] Weighting may not fully correct for existing differences between respondents and nonrespondents.

response rates and nonresponse adjustment factors for each nonresponse adjustment cell are shown in Exhibit E-1 for pilot and non-pilot establishments participating in the mail survey and in Exhibit E-2 for establishments participating in the on-site survey.

A third nonresponse adjustment weight was calculated for analyses that included establishments responding to both the mail and on-site surveys. This weight was based on establishments in the on-site sample. Only establishments that responded to both the mail and on-site surveys were considered respondents in calculating these rates. The appropriate adjusted weights were used in computing the means and percentages presented in this report.

Exhibit E-1: Response Rates for Pilot and Non-pilot Establishments Participating in the Mail Survey, by Size

Establishment Type	Size (employees)	Total	Percent of Non-pilot Nonrespondents	Percent of Non-pilot Respondents	Percent of Non-pilot Establishments	Response Rate (%)
Non-pilot	Unknown	35	1	14	5	83
	1-4	90	14	7	12	17
	5-9	94	14	9	13	21
	10-19	106	16	11	14	22
	20-49	98	13	14	13	31
	50-99	84	10	14	11	35
	100-249	96	14	11	13	25
	500-999	71	9	10	10	30
	1,000-4,999	43	6	6	6	30
	5,000-9,999	28	4	4	4	29
	Total	**745**				**28**

			Percent of Pilot Nonrespondents	Percent of Pilot Respondents	Percent of Pilot Establishments	
Pilot	Unknown	362	34	27	30	47
	1-4	100	10	6	8	41
	5-9	94	10	6	8	38
	10-19	45	4	3	4	47
	20-49	107	10	8	9	50
	50-99	88	7	8	7	58
	100-249	184	12	18	15	64
	500-999	94	8	8	8	54
	1,000-4,999	53	3	6	4	72
	5,000-9,999	72	2	9	6	81
	Total	**1,199**				**53**

Exhibit E-2: Response Rates for Pilot and Non-pilot Establishments Participating in the On-Site Survey, by Size

Establishment Type	Size (employees)	Total	Percent of Non-pilot Nonrespondents	Percent of Non-pilot Respondents	Percent of Non-pilot Establishments	Response Rate (%)
Non-pilot	Unknown	20	7	13	10	60
	1-4	22	13	9	11	36
	5-9	27	17	9	14	30
	10-19	23	13	10	12	39
	20-49	26	11	16	13	54
	50-99	28	16	12	14	39
	100-249	24	13	11	12	42
	500-999	18	6	12	9	61
	1,000-4,999	10	4	7	5	60
	5,000-9,999	1	0	1	1	100
	Total	**199**				**45**

			Percent of Pilot Nonrespondents	Percent of Pilot Respondents	Percent of Pilot Establishments	
Pilot	Unknown	101	44	27	29	84
	1-4	28	3	9	8	96
	5-9	33	8	9	9	91
	10-19	13	0	4	4	100
	20-49	22	6	6	6	91
	50-99	30	6	9	9	93
	100-249	74	17	22	21	92
	500-999	25	17	6	7	76
	1,000-4,999	16	0	5	5	100
	5,000-9,999	10	0	3	3	100
	Total	**352**				**90**

b. EMPLOYEE INTERVIEW WEIGHTS

The sample for the employee interviews was a stratified random sample, with differential sampling rates within strata to achieve oversampling of certain employee groups of special interest to the evaluation. The sampling strata were defined based on the agency confirming work authorization (SSA or INS) and the verification outcome from the Basic Pilot transaction database. Because equal probability sampling was not used in the employee interviews, the base weights were computed as the reciprocal of the probability of selection in the stratum, as shown in the last column of Exhibit E-3.

Exhibit E-3: Sample Size and Initial Weight by Employee Sampling Stratum

Sampling Stratum	Definition	Sample Size	Percent of Total Sample	Sampling Weight*
1	Initially authorized – INS	600	13	8.9367
2	Initially authorized – SSA	800	17	50.0325
3	Authorized 2nd stage – INS	763	16	3.4988
4	Eventually authorized – SSA	800	17	14.9113
5	Authorized 3rd stage – INS	142	2	1.0000
6	Self-terminated/quit – INS	348	8	1.0000
7	Self-terminated/quit – SSA	350	8	12.7086
8	Unconfirmed – SSA	350	8	7.0171
9	Unauthorized – INS	114	2	1.0000
10	Inconsistent authorization outcome – INS	443	9	1.0000
	Total	**4,710**		

* A weight of 1 indicates that all in-scope verifications in the stratum were selected.

As anticipated, the results of tracing, locating, and interviewing the original sample of 4,710 employees differed by sampling stratum. To reflect these differences, the sampling weights were adjusted for nonresponse independently within sampling stratum. After reviewing the nonresponse patterns, the evaluation team defined nonresponse adjustment cells to reflect different nonresponse patterns by sampling stratum and place of birth (native-born and foreign-born). Exhibit E-4 shows the definition of the adjustment cell and the value of the weights adjusted for nonresponse. These adjusted weights inflated the employee sample data to the population represented by the sample, namely, employees verified by pilot establishments during the study reference period. These adjusted weights were used in computing the means and percentages presented in this report.

Exhibit E-4: Employee Final Weights and Nonresponse Adjustment

Stratum	Sampling Stratum	Substratum	Original Base Weight	Nonresponse Adjustment	Final Weight
1	Initially authorized – INS	Foreign-born	8.9367	3.39	30.2938
2	Initially authorized – SSA	Foreign-born	50.0325	4.66	232.9099
		Native-born	50.0325	4.43	221.8108
3	Authorized 2nd stage – INS	Foreign-born	3.4988	3.71	12.9759
4	Eventually authorized – SSA	Foreign-born	14.9113	4.75	70.8841
		Native-born	14.9113	3.13	46.7410
5	Authorized 3rd stage – INS	All	1.0000	7.10	7.1000
6	Self-terminated/quit – INS	All	1.0000	5.12	5.1176
7	Self-terminated/quit – SSA	All	12.7086	10.61	134.7879
8	Unconfirmed – SSA	All	7.0171	5.00	35.0857
9	Unauthorized – INS	All	1.0000	7.60	7.6000
10	Not categorized	All	1.0000	6.42	6.4203

3. VARIANCE AND TESTS OF SIGNIFICANCE

In sample surveys, the variance is used to estimate the precision of the survey estimates. This is especially important when the researcher wants to know whether any differences observed can be explained by chance. Tests of significance indicate the likelihood that observed differences occurred by chance.

Statistical software packages such as SPSS and SAS permit variance estimation for equal probability samples when there is no nonresponse weighting adjustment. These variance estimates are not appropriate for use with the sampling and nonresponse adjustment methods used for the employer survey and the employee interviews. Instead, the evaluation team used WesVarPC for variance estimation, because this software accounts for differential weighting and nonresponse adjustment among respondent subgroups in the computation of test statistics.

The variance estimates generated by WesVarPC were used for statistical tests of significance. The significance level for the statistical tests was set at 0.05 ($\alpha = 5\%$), a commonly used significance level.

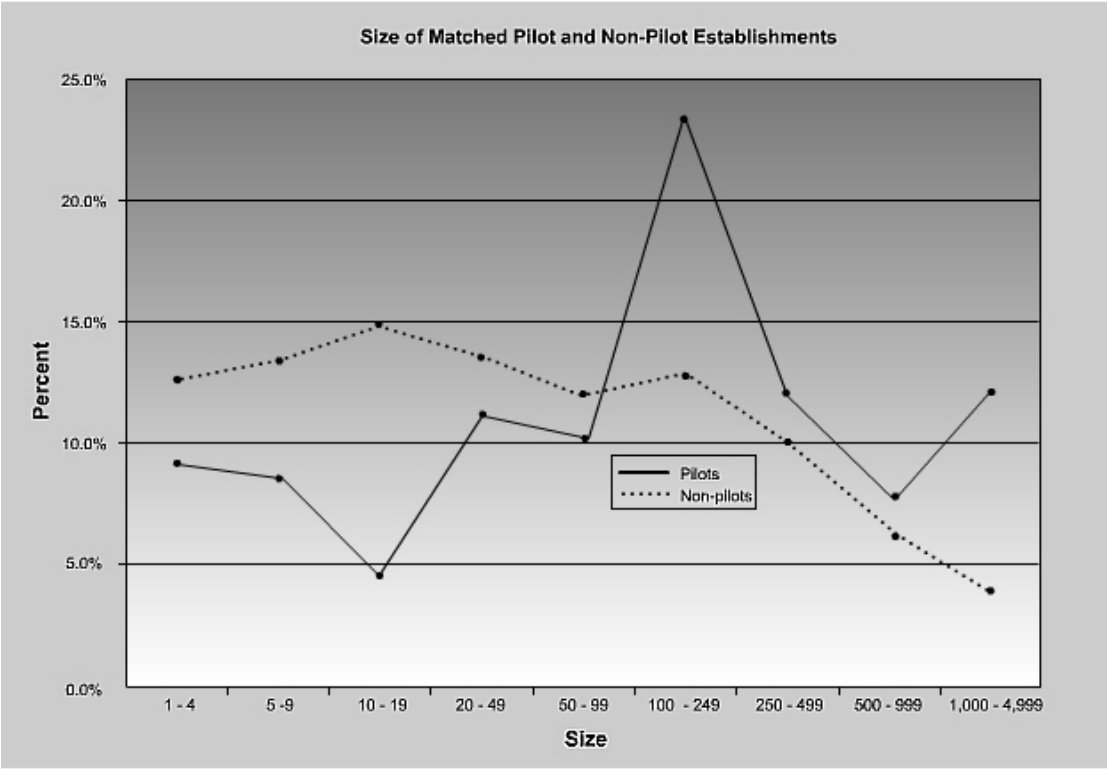
B. ESTABLISHMENT MATCHING

The establishments that served as candidates in matching non-pilot to active pilot establishments were obtained from the GENESYS business database developed by Marketing Systems Group (MSG). The matching to pilot employers was based on the establishment's county of physical location, four-digit SIC code, and employee-size class. In each county-SIC-size cell, MSG selected up to five possible matching candidates. If there were no matching candidates in the county-SIC-size cell, the size cell boundary was relaxed until a match was found.

To determine the overall success of the matches, the evaluation team compared the aggregate distribution of matched pilot and non-pilot establishments on the three matching variables. Exhibit E-5 shows that the match on establishment size was not completely successful. Compared to pilot establishments, non-pilot establishments tended to be more concentrated in the under-100 employee-size categories and less concentrated in the higher categories.[2]

[2] This disparity presumably occurs because of large size differences between pilot and non-pilot establishments in the total population. For example, only 0.5 percent of the establishments on the GENESYS database have 250 or more employees, while 32 percent of pilot establishments were in this size category. These size differences make it difficult to match large pilot establishments on size, county, and four-digit SIC code. As a result, when the matching criteria are relaxed, non-pilot establishments are likely to be consistently smaller than the pilot establishments to which they have been matched.

Exhibit E-5: Size of Matched Pilot and Non-pilot Establishments



SOURCE: GENESYS Database, Marketing Systems Group

Since the match was not entirely successful, it is not possible to make inferences about the likely effects of the pilot program through a simple comparison of pilot and non-pilot employers. Therefore, multivariate regression techniques were used to "control" for the remaining differences between pilot and non-pilot employers to examine whether there is more discrimination among pilot employers than among non-pilot employers.

C. FEDERAL GOVERNMENT COST ESTIMATES

Chapter X provides basic information about the assumptions underlying the cost estimates in Chapters X and XII. This section provides additional details about the assumptions.

1. STEPS UNDERTAKEN TO OBTAIN COST ESTIMATES FOR THE BASIC PILOT PROGRAM

a. OBTAINING PRELIMINARY ESTIMATES

The first step in calculating Federal estimates was to obtain from INS staff estimated total costs for the Basic Pilot program. These preliminary estimates covered the period from January 1997 through April 2000. As noted in Chapter X, no formal records were available for estimating many of these costs.

b. Breakdown of Current Costs into Start-up Costs Versus Operating Costs

Evaluation staff made a preliminary attempt to divide the costs provided by INS into start-up and operating costs. They then met with INS staff to review these estimates, clarify what information was included in the various categories, and determine how costs should be prorated across programs. For example, some of INS's original cost estimates were for all Systematic Alien Verification for Entitlements (SAVE) programs rather than just costs associated with the Basic Pilot program. Fixed costs that would have been incurred in the absence of the pilot programs (e.g., development costs for software used by SAVE to administer both the benefit programs and the pilot programs) were excluded from the cost figures.[3] Other costs (e.g., Headquarters salaries for pilot program staff) were prorated among the pilot programs.

c. Further Breakdown of Operating Costs

After estimating a total operating cost for each major item, the evaluation team annualized the operating costs by dividing the historical costs by the number of years on which they were based. No adjustment was made for inflation.

For each of the major cost categories listed in Exhibit X-1, the evaluation team broke the estimated annual cost into the following types of costs[4]:

- Fixed costs
- Cost per establishment
- Cost per new establishment
- Cost per SSA office
- Cost per INS office
- Cost per query
- Cost per SSA referral visit
- Cost per INS second-stage verification

For each cost category, the evaluation team also estimated average annual number of units.

---

[3] Excluded were costs for developing the Verification Information System (VIS) ($6.25 million) and the Web Access Method of Employment Verification ($510,000). VIS and Web Access comprise the new equipment and software platform developed for all of the SAVE programs.

[4] For any major cost category, only some of the types of costs were applicable. The relevant types of costs for each major category are indicated in Exhibits X-2 through X-7.

2. ADDITIONAL NOTES ON SPECIFIC COSTS

a. INS HEADQUARTERS SALARIES AND EXPENSES

INS was able to provide a good estimate for the total salaries of SAVE staff working on the pilot employment verification projects between January 1997 and April 2000. SAVE staff estimated that approximately 80 percent of their time was devoted to the Basic Pilot program. This high estimate sounded reasonable, given that the Basic Pilot program was by far the largest pilot program operating during this time. As noted in Chapter X, the primary pilot-related tasks of INS Headquarters staff are to develop policies and systems for the pilot programs. These functions should not be particularly sensitive to measures of the size of the pilot program, such as the number of employers. Therefore, the evaluation team estimated that 85 percent of INS Headquarters annual pilot-related costs were for fixed expenses.

Because of the accounting methods used at INS, individual operating units are aware of their total salary costs but not all of the non-salary expenses associated with their employees. Many indirect costs, such as employee benefits, are handled by other parts of the agency. INS provided an internal report that broke down total costs associated with different types of employees into salary, overhead items (such as training, communications equipment, and office supplies), and employee benefits. An analysis of this information indicated that total INS costs per employee were approximately 2.5 times employee base salaries. Therefore, the evaluation team estimated that the total salary and expenditure costs for the pilot program were 2.5 times the INS estimate of salary costs.

INS staff also estimated the percentage of their time spent in dealing with establishments, but they could not estimate how their time was split between new establishments and those that had been in operation for a longer period. The evaluation team believed that establishments new to the pilot program would have more questions than those with more experience of the program; therefore, the total expenditure for answering questions was broken into costs per new employer and costs per existing employer.

b. INS FIELD PERSONNEL

INS provided the evaluation team with the number and approximate grade levels of Immigration Status Verifiers (ISVs) and their supervisors employed in the Los Angeles office.[5] To estimate the total salary expense for these employees, the evaluation team used Federal salary levels for Los Angeles in 2000 for employees at step 3 in their respective grades.

The ISVs and their supervisors both indicated that the normal workload for an ISV was 70 cases a day. The average ISV was assumed to work 218 days a year (i.e., 261 total

[5] INS indicated that the ISVs were at either the GS-5 or GS-7 level but did not indicate what percentage of staff were at these levels. Therefore, the evaluation team used a salary level halfway between the GS-6 and GS-7 levels. Supervisors were at the GS-9 level.

weekdays a year minus 13 sick days, 20 vacation days, and 10 holidays). Thus, the average ISV was assumed to complete 15,260 cases a year. Since there were 3 supervisors for 21 ISVs[6] in the Los Angeles office, the salary costs for these 70 cases were assumed to include one-seventh of the cost of a GS-9 in addition to the salary cost of the ISV. For field staff, the evaluation team used the same ratio of total employee-related costs to salaries (2.5) that was used for Headquarters staff. The total estimate of employee costs for the Los Angeles ISVs was $5.62 per case, calculated as shown below:

$$[(\$28,770 + \$38,945/7)*2.5]/15,260$$

The estimate of $5.62 does not include costs associated with ISVs in other offices. The Los Angeles ISVs indicated that these other ISVs were involved only in "walk-in" cases in their own offices (a rare occurrence) and in cases where the Los Angeles ISV needed information from the noncitizen's local office. The evaluation team believed that it was reasonable to incorporate the costs for these ISVs by adding 10 percent to the employee costs for the Los Angeles ISVs, for a total cost of $6.19 for cases that go to secondary verification at INS.[7]

c. SSA Salaries and Expenses

INS reimburses SSA for any costs incurred in administering the Basic Pilot program. According to INS staff, SSA billed INS $1.8 million for Basic Pilot costs for the period of January 1997 to April 2000. However, SSA staff indicated that the billed amounts were rough estimates. For example, SSA's last bill was based on the assumption that 15 percent of cases were referred to SSA field offices. However, information from the Basic Pilot transaction database (presented in Chapter V) indicated that the correct rate was approximately 3 percent. Given this degree of inaccuracy in the SSA estimates, the evaluation team decided to estimate SSA costs on the basis of discussions with SSA staff (including a site visit to a Los Angeles office near several pilot employers), as well as information on comparable INS costs.

The site visit to the SSA field office indicated that SSA cases were no more complex than INS cases. Therefore, the cost for SSA cases was assumed to be the same as that estimated for INS cases (approximately $6).[8]

---

[6] The total salaries for these staff could not be used, because the ISVs also handled verification for benefit cases and a State-run pilot program, in addition to the INS employment verification programs.

[7] This cost includes any work that the ISVs must do for cases that go to the third stage, because Los Angeles ISVs are responsible for following cases to their conclusion whether or not they include a third-stage verification.

[8] This per-case cost may be an overestimate, since it does not include a correction for cases that would have come to SSA's attention even in the absence of the Basic Pilot program.

d. VERIFICATION QUERY COSTS

Because a contractor performed most of the verification query work, INS had good estimates of these costs.

e. PRODUCTION OF COMPUTER DISKS, MANUALS, AND PILOT NOTICES

Tasks such as producing computer disks, manuals, and notices typically entail a fixed cost for setting up the production process, plus a per-item cost. The breakdown of the costs into these components is approximate.

f. SYSTEMS OPERATION AND MAINTENANCE

INS reported that systems operation and maintenance costs depended on the number of problems encountered. The evaluation team assumed that most problems would arise regardless of the size of the system. However, it also seems reasonable that the likelihood of problems occurring (or being detected) increases with the number of system users. Therefore, the evaluation team assumed that some of the systems costs would be proportional to the number of establishments enrolled in the pilot program. The proration of fixed and variable costs was approximate.

3. ESTIMATION OF RELEVANT UNITS

a. ESTABLISHMENTS

Estimates of the total number of establishments in the Nation and within specified size categories were obtained from the Census Bureau's Web site (http://www.census.gov/epcd/www/smallbus.html). This site also listed information on the number of employees working for these establishments.

The number of establishments in the Basic Pilot program was calculated from the information INS maintained on the number of signed Memorandums of Understanding (MOUs) on a monthly basis. These monthly figures were cumulated to estimate the average number of employers with signed MOUs during the period for which the evaluation team had cost figures.

According to Census Bureau figures, 35 percent of the Nation's establishments in 1995 were in the five original pilot States.[9] Since some establishments outside of these States were also eligible for the Basic Pilot program, and since the original States had relatively large illegal immigrant populations, the evaluation team estimated that twice as many establishments would join a voluntary national program as had participated in the original States. This assumption is, of course, an arbitrary one.

[9] Office of Advocacy, U.S. Small Business Administration, from data provided by the U.S. Census Bureau (http://www.sbaonline.sba.gov/advo/stats/).

The enhanced Basic Pilot program, which should be more attractive to employers than the current pilot program, was also assumed to have twice the number of establishments as the Basic Pilot program. This is an arbitrary assumption.

b. NEW ESTABLISHMENTS

The number of establishments newly enrolled in the current Basic Pilot program was estimated by annualizing the number of monthly enrollments recorded by INS during the base period. For the other programs, the number of new establishments was estimated by multiplying the national ratio of new establishments to the total number of establishments (11.9 percent)[10] by the number of establishments estimated, as explained in the preceding section.

c. SSA OFFICES

The total number of SSA field offices (1,340) was obtained from the *Social Security Bulletin*, *Annual Statistical Supplement, 2000*.[11] All of the pilot programs except the current program were assumed to involve all of the SSA offices. Since the current Basic Pilot program primarily affects the five original States, the evaluation team assumed, somewhat arbitrarily, that only 300 SSA offices are currently involved.

d. INS OFFICES

INS provided a list of ISVs within INS field offices. According to this list, 50 of the 73 field offices currently have ISVs. The evaluation team assumed that all 73 INS offices would have responsibilities for any of the programs other than the current program.

e. QUERIES

The number of queries for the Basic Pilot program was obtained from the Basic Pilot transaction database. The voluntary national program and the enhanced Basic Pilot program were assumed to have twice the number of queries as the Basic Pilot program. For the mandatory programs, the number of queries was assumed to be proportional to the number of employees in the relevant employee-size class compared to the number of employees reported by Basic Pilot establishments.[12]

---

[10] Office of Advocacy, U.S. Small Business Administration, from data provided by the U.S. Census Bureau (http://www.sbaonline.sba.gov/advo/stats/).

[11] http://www.ssa.gov/statistics/Supplement/2000/admin.pdf.

[12] http://www.census.gov/epcd/www/smallbus.html.

f. SSA REFERRAL VISITS

The annual number of employees going to SSA field offices because of the Basic Pilot program was estimated from the Basic Pilot transaction database. The number of visits to SSA was assumed to be twice as high for the voluntary national program and the enhanced Basic Pilot as for the Basic Pilot program. For the mandatory programs, the number of SSA visits was assumed to be proportional to the number of employees in the relevant employee-size classes compared to the number of employees reported by Basic Pilot establishments.

g. INS SECOND-STAGE VERIFICATIONS

INS provided information on the number of Basic Pilot cases that were referred for second-stage verification. The evaluation team assumed that twice as many INS second-stage verifications would be required for the voluntary national program and the enhanced Basic Pilot program as for the Basic Pilot program. For the mandatory programs, the number of INS second-stage verifications was assumed to be proportional to the number of employees in the relevant employee-size classes compared to the number of employees reported by Basic Pilot establishments.

D. ESTIMATION OF THE WORK-AUTHORIZATION STATUS OF UNRESOLVED CASES

1. BACKGROUND

One of the evaluation's assigned tasks was to estimate the number of persons without work authorization who were actually deterred from working at establishments participating in the Basic Pilot. However, only 0.04 percent of all employees were actually determined by the Basic Pilot system to be unauthorized, while 13 percent of employees received tentative nonconfirmations that were never resolved. This unresolved group undoubtedly included many non-work-authorized individuals who opted not to contest the system's finding because they knew they would not be successful. However, the 13 percent of cases resulting in tentative nonconfirmations also includes employees who were work-authorized. In some of these cases, a work-authorized employee was never notified of the tentative nonconfirmation and was therefore unable to resolve it. In other cases, the worker may have decided that it wasn't worth the effort to resolve the case. This latter situation was especially likely, because jobs were plentiful during the data collection period and many of the jobs in question were disagreeable, high-turnover positions. The fact that some employers take adverse actions against employees receiving a tentative nonconfirmation further increases the probability that the employee will not contest the finding.

An analysis of 95 employees with unresolved tentative nonconfirmations indicated that this group includes substantial numbers of both work-authorized and non-work-authorized employees. One possible way of estimating the number of unauthorized workers in the unresolved category is to use the results of this small-scale study to estimate the total percentage of non-work-authorized cases. Applying the observed

percentage of resolved cases from this small-scale study to the total pool of unresolved cases, the evaluation team estimated that 5 percent of all Basic Pilot cases were non-work-authorized employees whose cases were never resolved. This estimate can be challenged, because the sample of cases is not representative of all unresolved cases. Most importantly, the sample included no cases in which the tentative nonconfirmation was issued by SSA, even though 91 percent of the unresolved cases were not referred to INS.

The evaluation team therefore developed a model that estimates the desired percentage, using information from the transaction database as a starting point. Since the processes for INS and SSA case are slightly different, they are examined separately below.

2. ESTIMATES FOR SSA CASES

Exhibit E-6 provides the basic model for cases in which SSA issues a tentative nonconfirmation that should be resolved with SSA. The bold letters in parentheses on the exhibit are for cross-reference purposes.

The purpose of the model is to estimate how many employees receiving final nonconfirmation outcomes from SSA would have been found to be work-authorized, given what is known about the cases and a set of "reasonable assumptions." For each SSA tentative nonconfirmation case, the transaction database provides the following information:

- Final case disposition (i.e., authorized or final nonconfirmation)
- Initial SSA finding (i.e., why the computer match failed)

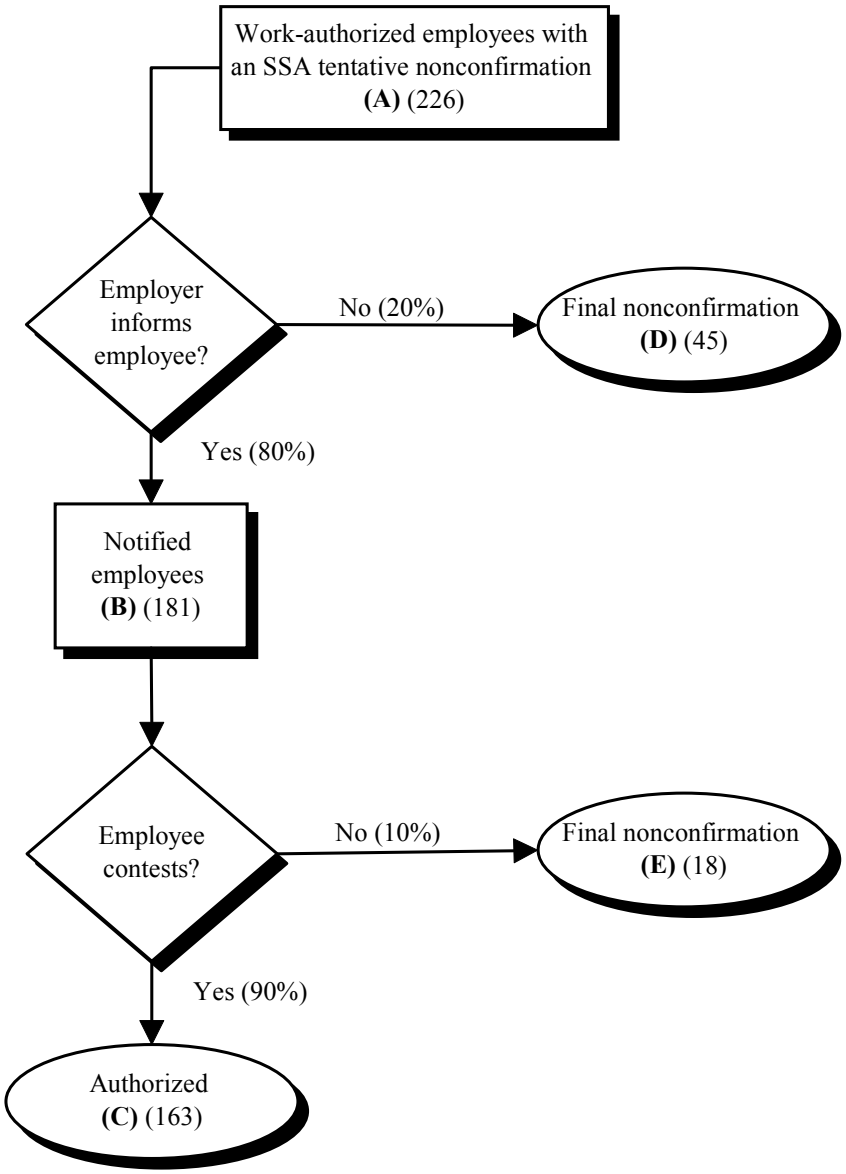
This information is insufficient for estimating how many work-authorized employees are in the final nonconfirmation category. Therefore, the evaluation team had to use additional reasonable assumptions to estimate the number of employees with final nonconfirmations who would have been determined to be work-authorized by the Basic Pilot system if all cases had been resolved. The following assumptions were used:

(1) The probability that an employee receiving a final nonconfirmation from SSA is actually work-authorized[13] depends on the initial disposition of the case. For example, there are likely to be more work-authorized individuals among those rejected because the employee's name did not match the SSA database than among those for whom both the name and date of birth disagreed. The reasonableness of this assumption is consistent with data on the percentage of employees in each category who contested tentative nonconfirmations. Work-authorized employees are presumably more likely than non-work-authorized employees to contest nonconfirmations. In fact, tentative nonconfirmations were

[13] To simplify the explanation of the model, employees who were or would have been authorized by the system are referred to as work-authorized. In reality, as discussed in the report, some employees determined to be work-authorized were not actually work-authorized.

contested by 66 percent of work-authorized employees, compared to 17 percent of non-work-authorized employees (Exhibit E-7). This analysis required that the evaluation team separately estimate the number of work-authorized employees within each SSA initial category. Exhibit E-6 depicts work-authorized employees rejected by SSA during the automated procedures because they had an invalid Social Security number.

Exhibit E-6: Basic Pilot Process Between SSA Tentative Nonconfirmation and Final Determination for Work-Authorized Employees*



* The numbers refer to employees with an initial SSA determination of "invalid Social Security number" and assume that 80 percent of employees are informed of the tentative nonconfirmation and that 90 percent of work-authorized employees contest the tentative nonconfirmation.

SOURCE: The number authorized (C) is from the transaction database. The remaining numbers are estimated.

Exhibit E-7: Percentage of Employees Receiving Tentative Nonconfirmations from SSA Whose Cases Were Resolved, by Initial SSA Finding*

Initial SSA Finding	Percent of Tentative Nonconfirmation Cases Resolved	No. of Tentative Nonconfirmations
Invalid Social Security number	2.8	5,644
Date of birth disagrees with SSA database	26.4	9,870
Name disagrees with SSA database	65.9	9,497
Name and date of birth disagree with SSA database	17.3	16,551
Social Security number belongs to deceased person	6.3	256
Unlawful permanent resident – INS must confirm work authorization	15.1	10,304
Facility unavailable	86.6	142
SSA record indicates that employee is foreign-born, but there is no citizenship code	0.4	1,960
Missing initial Social Security Code	0.3	2,742
Total		**312,640**

SOURCE: Transaction Database

* The table includes only cases that were not sent to INS.

(2) The percentage of employees informed by their employers of a tentative nonconfirmation from SSA does not depend on the reason for issuing the tentative nonconfirmation. For example, employees not matched because of an invalid Social Security number were no more or less likely to have been informed of nonconfirmation than were employees whose names did not match the SSA database. In the model, the user estimates this percentage, so alternative scenarios can be tested. Exhibit E-6 models the probability that the employee will move from **(A)** to **(B)** (from tentative nonconfirmation to notification). The illustration assumes that the user has set the percentage of notified employees equal to 80 percent.

(3) The percentage of work-authorized employees contesting SSA tentative nonconfirmations does not depend on the reason for issuing the tentative nonconfirmation. For example, employees not matched because of an invalid Social Security number are no more or less likely to contest than are employees who did not match on date of birth. In the model, the user estimates this percentage, so alternative scenarios can be tested. Exhibit E-6 models the probability that the employee will move from **(B)** to **(C)** (from notification to authorization). The illustration assumes that the user has set the percentage of employees who contest tentative nonconfirmations equal to 90 percent.

The number of employees determined to be work-authorized for a given SSA initial finding is known. For example, 163 employees with initially invalid Social Security numbers were authorized **(C)**.

Mathematically, the number of authorized cases **(C)** is equal to the number of contested cases **(B)** multiplied by the probability that a work-authorized employee will contest the finding. The probability that a work-authorized employee will contest a finding is a user-driven input. For illustrative purposes, assume that the model user has estimated that 90 percent of work-authorized employees will contest an SSA tentative nonconfirmation. In this case, **(C)** = 0.9 * **(B)**. Mathematically, this is equivalent to **(B)** = **(C)**/0.9. In the example, **(B)** = 163/0.9 = 181.[14]

The number of employees who were informed of a tentative nonconfirmation can be estimated in a similar fashion, given the user-input assumption of the percentage of work-authorized employees with tentative nonconfirmations who are informed of their status. Continuing with the example and assuming that 80 percent of employees with SSA tentative nonconfirmations are notified, the number of work-authorized employees can be estimated as **(A)** = **(B)**/0.9 = 226. Thus, the model estimates that the Basic Pilot would have found 226 employees with initially invalid Social Security numbers to be authorized, if all employees had been properly notified and had chosen to contest their cases.

3. INS PROCESS

For employees receiving tentative nonconfirmations from INS, information is available for initial INS status as well as initial SSA status. The INS model, therefore, classifies a case according to a combination of its INS status and its SSA status. Since some of the resulting combination cells had too few cases to be useful, small cells were combined. Except for this starting difference, the model for INS-issued tentative nonconfirmations (not shown) is directly parallel to that for SSA. The two user-entered variables (percentage of employees notified and percentage of notified employees who contest) are entered separately for INS and SSA. Separate variables are entered because INS and SSA cases are likely to be different, since employees with INS tentative nonconfirmations are noncitizens while those with SSA tentative nonconfirmations are citizens. Further, employers and employees know whether the tentative nonconfirmation is from INS or SSA.

4. ESTIMATED RANGE

To obtain a preliminary estimate of the range of possible values for the percentage of non-work-authorized employees, the evaluation team tested two scenarios. In the first, all of the user-input parameters were set to 100 percent. This scenario assumes that all work-authorized employees have already been identified by the system. As expected, with this approach the percentage of non-work-authorized employees is estimated at 13 percent. This is the maximum value.

[14] This means that (F) contains 0.2 * 195 = 39. This is not shown in Exhibit E-7, because this number was not directly used in the evaluation team's calculations.

To obtain a reasonable minimum value, the evaluation team assumed that 90 percent of employees receiving SSA tentative nonconfirmations were informed and that 75 percent of work-authorized employees who were informed contested the findings. Under the assumptions of the model, rates lower than these would be inconsistent with the observed rates of resolved cases for one or more of the SSA and INS categories respectively. In this situation, the estimated percentage of non-work-authorized employees is 10 percent. This range of 10 to 13 percent is considerably higher than the estimate from the special sample of 95 INS cases examined. It is interesting to note that, in this context, the model shows a higher percentage of unauthorized employees among individuals receiving SSA nonconfirmations than among those receiving INS nonconfirmations.